2cc2 4/23/03 #21 4/28/03 2700

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
Re: Appeal to the Board of Patent Appeals and Interferences

In re PATENT application of
MILLS

Group Art Unit: 2624

Application No. : 09/220,970

Examiner: Chen, W.

Filed: 12/23/98

Date: September 3, 2002

(September 1, 2002 = Sunday
September 2, 2002 = Holiday)

Hon. Asst. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

1 [ ] **NOTICE OF APPEAL**: Applicant hereby appeals to the Board of Patent Appeals and Interferences

2 [ ] **BRIEF** on appeal in this application attached in triplicate.

3 [X] An **ORAL HEARING** is respectfully requested under Rule 194 (due two months after Examiner's Answer -- unextendable).

4 [X] Reply Brief is attached in triplicate (due two months after Examiner's Answer -- unextendable).

5 [X] "Small entity" verified statement filed: [ ] herewith. [X] previously.

6 **FEE CALCULATION:**	Large/Small Entity	
If box 1 above is X'd, see box 12 below first and decide:... enter	$	$
If box 2 above is X'd, see box 12 below first and decide:... enter	$	$
If box 3 above is X'd, see box 12 below first and decide:... enter	$140	$140
If box 4 above is X'd, ... enter nothing	160	$ 160
7. **Original due date: September 1, 2002**		
8. **Petition is hereby made** to extend the original due date to cover the date this response is filed for which the requisite fee is attac (1 months) $ (2 months) $ (3 months) $ (4 months) $ (5 months) $		
9. Enter any previous extension fee paid [ ] previously since above original due date (item 7); [ ] with concurrently filed amendment........................	-	
10. **Subtract line 9 from line 8 and enter: Total Extension Fee**		
11. **TOTAL FEE ATTACHED =**		$ 300

12. [ ] *Fee NOT required if/since paid in prior appeal in which the Board of Patent Appeals and Interferences did not render a decision on the merits.

CHARGE STATEMENT: The Commissioner is hereby authorized to charge any fee specifically authorized hereafter, or any missing or insufficient fee(s) filed, or asserted to be filed, or which should have been filed herewith or concerning any paper filed hereafter, and which may be required under Rules 16-18 (missing or insufficient fee only) now or hereafter relative to this application and the resulting Official document under Rule 20, or credit any overpayment, to our Account/Order Nos. 50-0687/ 62-231 for which purpose a duplicate copy of this sheet is attached. **This CHARGE STATEMENT does not authorize charge of the issue fee until/unless an issue fee transmittal form is filed.**

Manelli, Denison & Selter, PLLC

Customer No.: 20736

Att: ______________
Jeffrey S. Melcher
Reg. No. 35,950
Tel: 202.261.1045
Fax: 202.887.0336

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re PATENT Application of
Mills

Group Art Unit: 2624

Application Ser. No. 09/220,970

Examiner: W. Chen

Filed: 12/23/98

For: A METHOD AND SYSTEM FOR PATTERN RECOGNITION AND PROCESSING

* * * * *

September 3, 2002
(September 1, 2002 = Sunday, September 2, 2002 = Holiday)

REPLY BRIEF

Hon. Asst. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

This Reply Brief to the Examiner's Answer dated July 1, 2002 is submitted in triplicate in accordance with 37 C.F.R. § 1.193(b) and directly responds to each new argument raised on pages 17-31 of the Examiner's Answer.

Claims 51-156, 160-265, 270-284, 287, 290-298, 300, and 304-312 all recite the use of a "Fourier series in Fourier space," a claim term that was coined by Appellant to distinguish over conventional Fourier series. The Examiner in his Answer, however, completely ignores this claim language to force fit his misapplication of prior art that is incapable of functioning in a way even remotely similar to the claimed invention. Recognition of these distinctions, among others, between Appellant's claimed invention and the prior art is what led six previous Examiners to allow the present application

before other PTO personnel suspiciously withdrew that allowance and transferred the application.

For example, in reference to claim 51, when constructing a Fourier series in Fourier space by parameterization with the data, the variable is frequency and the data parameters are constants. Each component is mathematically independent of another. This novel use a Fourier series in Fourier space cannot possibly read on conventional Fourier series, as alleged by the Examiner on pages 19-21 of the Examiner's Answer. A conventional Fourier series is based in time or x,y,z space, not Fourier space.

Furthermore, the conventional Fourier series is not parameterized with the data and each component has no meaning except in its entirety. For the sake of argument, if the individual components of the conventional Fourier series are modified by using the presently claimed steps as suggested by the Examiner, the result would be meaningless and have no relation to the real world. Conventional Fourier series must be taken as a whole. In stark contrast, each component of Applicant's Fourier series in Fourier space has a data parameter, and can optionally also have a tag, such as input context, general context, association context, and/or order format context, so that each component can be independently processed relative to the other components and still provide a relation to the real world. This is a fundamental and patentable difference between conventional Fourier series and Applicant's novel Fourier series in Fourier space.

The Examiner's misapplication of prior art is further illustrated by his exclusive reliance upon neural networks to reject the claimed invention. Appellant's invention is fundamentally different from neural networks that utilize only conventional Fourier series. Thus, such neural networks cannot possibly anticipate or make obvious claims reciting the use of "Fourier series in Fourier space" for the reasons just explained.

Moreover, neural networks function in a completely different way than the claimed invention as discussed below with specific reference to the claimed elements.

The Examiner's failure to recognize these and other fundamental differences between the claimed invention and prior art permeates his entire Answer and exposes the bankruptcy of the positions espoused therein.

A. Appellant's Invention is Unique and Represents a Significant Advancement in Pattern Recognition and Information Processing

The claim term "Fourier series in Fourier space" as used in claims 51-156, 160-265, 270-284, 287, 290-298, 300, and 304-312 distinguishes over conventional Fourier transforms that provide Fourier spectra as cited by the Examiner. The claimed data format provides for unique memory structure and processing steps, such as the determination of the spectral similarity as recited in claims 51-117, 119, 127-155, 160-227, 229, 237-265, 271-280, 282, 288, 290-293, 294-298, 301, 306 and 313-322. The claimed data format also allows for encoding context as a modulation of each Fourier component in Fourier space corresponding to delays as recited in Claims 65-86 and 175-196. The representations of physical objects with physical context as a Fourier series in Fourier space may also be performed by using an equivalent memory structure, wherein the memory formatted data is used directly during processing, such as in the determination of β_s^2 of claims 98-101, 106-109, 208-223, and 258-265, without the need for the step of constructing the specific Fourier series in Fourier space as recited in claims 61 and 171. None of these characteristics are present in the use of conventional Fourier series.

In claims 51-156, 160-265, 270-284, 287, 290-298, 300, and 304-312, each component representative of a characteristic of a physical object is <u>independent</u> of any other component. In contrast, each component of a conventional Fourier series has no meaning with regard to the representation of any real world object. Only the totality of

the components in a conventional Fourier series has any physical meaning, and no single component may be independently modified without losing the connection to the real world object that the total series represents. Thus, for instance, the method of order formatting of strings according to the method of claims 127-156, 237-266, 271-289, 294-312 and 315-318 cannot be reproduced using conventional Fourier series in neural networks. The ability to encode context of the ordered strings using modulation of the Fourier series at data parameterized frequencies, as recited in claim 266 also cannot be reproduced by using conventional Fourier series. In addition, the application of probability as the basis of forming associations and using probability based on prior activation rate as a basis to activate the components, as recited, for example, in claims 157-159, 267-269, 285-289, and 299-303, are unique.

An efficiency and universal applicability arise from the claimed structure. In contrast, the prior art processing based on standard Fourier series or transformed data stream and processing with neural networks has limited applicability. Based on the definitions of a standard transforms and neural networks and the corresponding mathematical representations and processing of data, the prior art can not perform the processing of a Fourier series in Fourier space or replicate the functions of the Appellant's invention, as discussed more fully below. Thus, all of the claims are allowable over the prior art.

B. Appellant's Invention of a Fourier Series in Fourier Space is Unique Over Traditional Transformed Data

As explained above, the "Fourier series in Fourier space" recited in claims 51-156, 160-265, 270-284, 287, 290-298, 300, and 304-312 does not read on conventional Fourier series. The claim term "Fourier series in Fourier space" is also distinguished from conventional Fourier series for the following additional reasons.

The prior art cited by the Examiner teaches the use of a conventional Fourier transform of the form given by equations on pages 19-21 of Examiner's Answer, with

subsequent processing using neural networks. If Appellant's data structure as provided for example, in claim 61, was equivalent to the conventional Fourier series cited by the Examiner, then the inverse Fourier transform must result in the original data stream as required by conventional Fourier series. However, it does not. In contrast, the inverse Fourier transform of the conventional Fourier series cited by the Examiner does result in the original data stream. Thus, Applicant's Fourier series in Fourier space cannot possibly be read on any prior art that teaches the exclusive use of a conventional Fourier transform.

A specific working example of the Fourier series in Fourier space is described in the present specification at page 8, line 19 to page 9, line 9 to provide guidance to one skilled in the art as to how to formulate Applicant's Fourier series in Fourier space. The Examiner has allowed claims reciting this specific example. Appellant, however, is entitled to the broader claims that cover other embodiments of Fourier series in Fourier space since they clearly distinguish over the cited prior art. This working example, which discloses a relationship between the data recorded from the physical world by transducers and the parameters of the Fourier series in Fourier space, merely illustrates this broader distinction:

> Referring to FIGURE 2, in the first step, the Input Layer 12 receives the data from the transducer (not shown). A Fourier transform processor 22 encodes each data element as parameters of a Fourier component in Fourier space and stores the data parameter values to the Input Layer section 24 of the memory 20. Each Fourier component of the Fourier series may comprise a quantized amplitude, frequency, and phase angle. For example the Fourier series in Fourier space may be:
>
> $$\sum_{m=1}^{M}\sum_{n=-\infty}^{\infty}\frac{4\pi}{1+\frac{k_z^2}{k_\rho^2}}a_{0_m}N_{m_{\rho_0}}N_{m_{z_0}}\sin\left(\left(k_\rho-n\frac{2\pi}{\rho_{0_m}}\right)\frac{N_{m_{\rho 0}}\rho_{0_m}}{2}\right)\sin\left(\left(k_z-n\frac{2\pi}{z_{0_m}}\right)\frac{N_{m_{z0}}z_{0_m}}{2}\right)$$
>
> having a quantized amplitude, frequency, and phase angle, wherein a_{0_m} is a constant, k_ρ and k_z are the frequency variables, n, m, and M are integers, and $N_{m_{\rho_0}}$, $N_{m_{z_0}}$, ρ_{0_m}, and z_{0_m} are the data parameters.

> In a first embodiment, the data parameters $N_{m_{\rho_0}}$ and $N_{m_{z_0}}$ of the Fourier series component are proportional to the rate of change of the physical characteristic. Each of the data parameters ρ_{0_m} and z_{0_m} of each Fourier component is inversely proportional to the amplitude of the physical characteristic. In the triangle example, the amplitude of the voltage at a given CCD element relative to the neighboring CCD element defines the rate of change of the voltage which is converted into the data parameters $N_{m_{\rho_0}}$ and $N_{m_{z_0}}$. The inverse of the amplitude of the voltage of each CCD element is converted into the data parameters ρ_{0_m} and z_{0_m}. As illustrated in FIGURE 3 and described above, for each CCD element, the Fourier series, parameterized accordingly, are stored to a specific sub register 27 of a specific register 26 of the Input Layer section 24 of the memory 20. Since the structure of a Fourier series is known in the art, only the parameters need to be stored in a digital embodiment.
>
> The data parameterization and use of the resulting Fourier series in Fourier space in effect "compresses" the possible continuous stream of physical characteristics from the world and allows for processing with much less data. Each also allows each "data element' comprising as Fourier component in Fourier space to be independent of any other component.

This processing efficiency aspect is further disclosed at page 36, line 37 through page 37, line 6 of the present specification:

> And, the number of terms necessary to represent most objects is not overwhelming. In fact, even a potentially challenging object having sharp edges such as a square pulse poses no difficulty in that is fairly accurately represented by only seven terms of a Fourier series in the time domain comprising the prior art [1]. The same principle applies to information represented as a Fourier series in k,ω -space.

One skilled in the art would easily recognize how to build a Fourier series parameterized according to data as disclosed wherein <u>each Fourier component has frequency as the variable rather than time of the prior art</u>. For example, Claim 61 recites that k_ρ and k_z are the frequency variables. The parameters derived from the data are substituted into the Fourier series formula. For example a_{0_m} is a constant, n, m, and M are integers, and $N_{m_{\rho_0}}$, $N_{m_{z_0}}$, ρ_{0_m}, and z_{0_m} are the data parameters as recited

by Claim 61. This procedure provides a Fourier series in Fourier space having a quantized amplitude, frequency, and phase angle.

Appellant defined in his claims this concept of a Fourier series parameterized with the data wherein each Fourier component has frequency as the variable instead of time by using the claim term "Fourier series in Fourier space." The Examiner completely ignores the highlighted phrase "in Fourier space" in his nonsensical interpretation of this claim term as reading on conventional "Fourier series":

> The Examiner is entitled to give the broadest interpretation of the language of the claims. The Examiner is not limited to Applicant's definition which is not specifically set forth in the claims. [Answer at pages 25-26 (emphasis added).]

Patent law principles require the Examiner to construe all of the claim terms in applying the prior art. The Examiner's strained interpretation of Appellant's claims, however, which reads the term "Fourier series in Fourier space" out of the claims, violates those principles and should therefore be overturned.

Furthermore, because the representations of the data of the present invention and those from conventional Fourier transforms are different, any filtering of this data structure must give a mathematically different structure. The same applies to modulating the representation of the data to encode context. Even if the context encoding methods and functions of the prior art were the same, which they are not, the final mathematical representations of the data and the relationship to the corresponding physical object must be different.

C. Neural Networks Cannot Reproduce the Function or Utility of Appellant's Claimed Invention

In his Answer, the Examiner erroneously alleges that processing by neural networks is equivalent to the processing performed by the Appellant's invention. The Examiner merely copies down the steps of the Appellant's claims and then attributes

them to prior art teachings of neural networks even though these steps are not disclosed in neural networks. The Examiner is apparently confused as to the nature of neural networks and the manner by which they process data.

The Examiner further ignores Appellant's express teaching against using neural networks, which are incapable of reproducing the data processing functions of the present invention. Thus, the present invention is not anticipated by, or obvious over, the neural networks disclosed in the prior art cited by the Examiner.

The distinctions between Appellant's claimed invention and neural networks are made clear at the outset in the "Background of the Invention" section of the present specification:

> Background Of the Invention
>
> Attempts have been made to create pattern recognition systems using programming and hardware. The state of the art includes neural nets. Neural nets typically comprise three layers--an input layer, a hidden layer, and an output layer. The hidden layer comprises a series of nodes which serve to perform a weighted sum of the input to form the output. Output for a given input is compared to the desired output, and a back projection of the errors is carried out on the hidden layer by changing the weighting factors at each node, and the process is reiterated until a tolerable result is obtained. The strategy of neural nets is analogous to the sum of least squares algorithms. These algorithms are adaptive to provide reasonable output to variations in input, but they can not create totally unanticipated useful output or discover associations between multiple inputs and outputs. Their usefulness to create novel conceptual content is limited; thus, advances in pattern recognition systems using neural nets is limited.

A succinct definition of a neural network is given by B. Muller et al. [B. Muller, J. Reinhardt, M. T. Strickland, Neural Networks An Introduction, 2nd Edition, Springer, Berlin, (1995)]. On page 13 of Muller appears

21.1 A Definition

Neural network models are algorithms for cognitive tasks, such as learning and optimization, which are in a loose sense based on concepts derived from research into the nature of the brain. In mathematical terms *a neural network model is defined as a directed graph with the following properties*:

1. A state variable n_i is associated with each node i.
2. A real-valued weight ω_{ik} is associated with each link (ik) between two nodes i and k.
3. A real-valued bias ϑ_i is associated with each node i..
4. A transfer function $f_i[n_k, \omega_{ik}, \vartheta_i, (k \neq i)]$ is defined, for each node i, which determines the state of the node as a function of its bias, of the weights of its incoming links, and of the states of the nodes connected to it by these links.

In standard terminology, the nodes are called *neurons*, the links are called *synapses*, and the bias is known as *activation threshold*. The transfer function usually takes the form $f\left(\sum_k \omega_{ik} n_k - \vartheta_i\right)$ where $f(x)$ is either a discontinuous step function or its smoothly increasing generalization known as a sigmoidal function. Nodes without links toward them are called *input* neurons ; *output* neurons are those with no link leading away from them. A *feed-forward* network is one whose topology admits no closed paths.

Based on this definition of a neural network, the formation of association given, for example, in Claim 51, cannot be achieved with standard transforms or a neural network. To illustrate this point, Appellant provides the following commentary comparing the steps of Claim 51 (shown in italics) to a neural network:

51. A method for recognizing a pattern in information comprising data, the method comprising:

inputting data;

encoding data as parameters of a plurality of Fourier components in Fourier space;

A standard transform is not the same as the plurality of Fourier components in Fourier space.

adding at least two of said Fourier components together to form at least one Fourier series in Fourier space;

A neural network comprises a plurality of nodes with inputs and output links. Node output to at least one link is based on its transfer function $f_i[n_k, \omega_{ik}, \vartheta_i, (k \neq i)]$. A neural network could not perform this step independently of the state of the entire network, and the output alters the network which subsequently alters the outcome of any subsequent output. In addition, the output must be assigned to the operation of addition of the Fourier components by the programmer since the neural network acts as a binary switch. This is against the teaching of the present invention of deterministically adding said input components.

sampling at least one of said Fourier series in Fourier space with a filter to form a sampled Fourier series;

modulating said sampled Fourier series in Fourier space with said filter to form a modulated Fourier series;

determining a spectral similarity between said modulated Fourier series and another Fourier series;

determining a probability expectation value based on said spectral similarity;

generating a probability operand based on said probability expectation value;

selecting a desired value for said probability operand, wherein recognition a pattern in said information is obtained when said probability having said desired value; and outputting a recognized pattern.

A neural network comprises a plurality of nodes with inputs and output links. The nodes of any given layer are not interconnected. The links only exist between layers. Node output to at least one link is based on its transfer function $f_i[n_k, \omega_{ik}, \vartheta_i, (k \neq i)]$. A

neural network could not directly perform any of the steps of Claim 51, since they are specific mathematical operations, not a binary decision in each case based on a transfer function, biases, weighted inputs, outputs as well as intrinsic and extrinsic node states. As given by Muller on page 61 [B. Muller, J. Reinhardt, M. T. Strickland, Neural Networks An Introduction, 2nd Edition, Springer, Berlin, (1995)] when considering the multiplication of two functions, "the synapses of neural networks are summation, not multiplication devices. In other words, signals can only be added, but never multiplied at a synapse". The specific steps taught by the present Invention must be independent of the input-output operation of a neural network. Thus, to say that neural networks can be used to replicate the present invention requires: 1) that the structure of the data representations be equivalent; 2) each specific mathematical operation be identical; and 3) the neural network always output the equivalent result from all binary decisions over all nodes which matches the function of the steps of Claim 51. This would require the trivial output of 1 except for the step of

> *selecting a desired value for said probability operand, wherein recognition a pattern in said information is obtained when said probability having said desired value; and outputting a recognized pattern.*

And, even in this case where the output according to Claim 51 maybe 0 or 1, for example, the determination is based on the output of a probability operand, which is patentably distinct over a neural network defined *supra*. The output according to the last step of Claim 51 is based on the probability expectation value and a selected probability distribution. In contrast, the output of a neural network depends on its history. The neural network output is based on transfer functions, biases, weighted inputs, and outputs as well as intrinsic and extrinsic node states that all change with history. Thus, it is patently clear that a neural network cannot function according to the claimed invention.

These same arguments equally apply to establishing an order formatted pattern in information with respect to standard ordered information as recited in claims 127-156,

237-266, 271-289, 294-312 and 315-318. In contrast, neural network learning involves adjustment of the *synaptic* strengths. In one case, the output of a neural network is compared to desired standard information and the error between the output and that desired is determined and used to adjust weighting functions in an attempt to reduce the error as discussed by Muller in Chap. 6.2 [B. Muller, J. Reinhardt, M. T. Strickland, Neural Networks An Introduction, 2nd Edition, Springer, Berlin, (1995)]. For example, a back projection of the errors is carried out by adjusting all parameters that give rise to the output, such as making synaptic adjustments, and the process is reiterated until a tolerable result is obtained.

Clearly, the specific steps of solving a problem by the methods of claims 127-156, 237-266, 271-289, 294-312 and 315-318 cannot be replicated by a neural network considering that the method by which a neural network solves a problem cannot be known. Muller at page 20 discloses the nebulous nature of neural networks:

> A major disadvantage of neural networks is the broad lack of understanding of how they actually solve a given cognitive task. Our present ignorance stems from the fact that neural networks do not break a problem down into its logical elements but rather solve it by a holistic approach, which is hard to penetrate logically. When the logical structure of the solution of the problem is not known anyway, this represents an advantage (see above), but it becomes a disadvantage when one wants to determine whether the solution provided by the neural network is correct. The sole presently known method of testing the operation of a neural network is to check its performance for individual test cases, a not very enlightening technique. No one knows how to judge the performance of a neural network knowing only its architecture, and it is almost impossible to determine what task the network actually performs from pure knowledge of the synaptic efficiencies.

Since it is impossible to know what steps a neural network performs, the teaching of the use of neural networks cannot anticipate or render obvious Appellant's claimed invention that explicitly recites what steps are performed.

In addition, activation of each *neuron* of a neural network is based on an algorithm using the state of each *neuron*, as well as the state of the entire neural network, including all links and nodes connected through those links. It is not based on the prior activation rate of each processing component independently of any other as is the case in Appellant's invention. Thus, for this additional reason, the present invention is patentably distinct over prior art systems that depend on neural networks.

D. Claims 157 and 266-267 are Not Anticipated by Kortge

The Examiner's reliance on Kortge to reject claims 157 and 266-267 as anticipated under 35 U.S.C. § 102(b) is misplaced. The use of probability by Kortge to determine the index of the most probable class does not anticipate using a probability operand to determine the state of activation of a component according to the present invention. Furthermore, the neural network of Kortge is not the same as a probability operand according to the present invention.

In Kortge's scheme, a pattern is input to the network and an index of the most probable class is the output 36 of Fig. 3 of Kortge. The use of probability to determine the index of the most probable class is not the same as the use of a probability operand as taught by the Appellant. The inputs and outputs are entirely different. The input according to claims 157 and 267 is a number like a probability expectation value, such as a number between zero and one, rather than a pattern. The output according to claims 157 and 267 is activation rather than pattern generation or index of the most probable class. Furthermore, a probability operand does not change. It is based on a specific probability distribution, such as Gaussian. In claims 157 and 162, the input is the activation probability parameter and output is activation or nonactivation based on the operand. The activation probability parameter does change based on prior activation probability parameter and a weighting based on an activation rate of the corresponding component. Thus, it is updated and is dynamic. Claims 157 and 267 concern the activation of the components of the invention. The disclosure of Kortge cannot teach activation of components based on probability since knowledge of which particular components are involved, their state of activation, and the method that a neural networks solves a problem are unknown and cannot be determined as discussed above.

The Examiner states that Claims 157 and 162 recite no feature that will exclude reading of the claims from any network. This is simply not true. Neural networks are dynamic, not static, and thus cannot be used as a probability operand, as shown by the definition given by Muller above. The Examiner's statement is off base.

On pages 24-26 of the Answer, the Examiner has distorted Kortge's disclosure at column 7, lines 18-24 by taking it out of context. Kortge discloses that the output is the probability that the input is a recognized pattern, "[f]or example the output signal 36 may represent degrees of membership in various classes, such as the probability a handwritten character input is "A", "B", "C", etc." [See column 7, lines 9-12 of Kortge.] In this case, the input is not a number representative of the past and current activation of a component. A probability operand is not taught. The action of Kortge is not activation of a component of the recognition system as in the current Invention. Kortge teaches at column 7, lines 20-24, "[f]or example, if the system is being used to recognize handwritten characters, the effector 38 might store an ASCII representation of the most probable character into computer memory, perhaps to allow a user to send email using a device too small for a keyboard."

E. Claims 271-272, 274, 276, 278, 281-283, 285-288, 290-291, 299-301, 304, 309, and 312-320 Are Not Anticipated by Caid

Similarly misplaced is the Examiner's reliance on Caid to reject claims 271-272, 274, 276, 278, 281-283, 285-288, 290-291, 299-301, 304, 309, and 312-320 as anticipated under 35 U.S.C. § 102(b). None of these claims are anticipated by Caid for the many reasons provided in Appellant's Brief and as follows.

Claims 271-272, 274, 276, 278, 281-283, 287, 290-291, 300, 304, 309, and 312-320 are not anticipated by Caid since the conventional transform used by Caid is distinguished from a "Fourier series in Fourier space," as discussed above and in Appellant's Brief.

Specially, Caid discloses a means to represent images for later retrieval by forming a vector that is representative of attributes of the image. The vector is generated by a wavelet transformation of the image data. This wavelet transformation to form a Fourier series, however, is fundamentally different from that taught by Appellant. Caid creates a Fourier series wherein time or space is the running variable—not frequency as is the case in Appellant's invention. Caid's method of processing is entirely different since it operates in a different space than the claimed invention and, thus, the Examiner's strained attempts to connect Caid's method to the claimed invention is meritless.

These differences are not insignificant and go the heart of what distinguishes Applicant's claimed invention from the prior art. For example, since Caid's wavelet transformation is not in Fourier space, context cannot be encoded as time delays as taught by Appellant. Furthermore, at column 5, lines 8-11, Caid teaches that: "Random high dimensional context vectors are assigned to each atom. The context vectors are assigned to each atom. The context vectors are then modified according to the spatial relationship and co-occurrence of the atoms in the images in a procedure called bootstrapping." Caid's method of encoding is ad hoc and far different from the claimed invention.

Caid also does not disclose the use of a "activation probability operand," or any of the multitude of steps recited in claims 285, 286, 288, 299 and 301.

For these many reasons the Section 102 rejection over Caid should be withdrawn.

F. Claims 158-159 and 268-269 Are Not Obvious Over Kortge as Applied to Claims 157 and 267 Above, and Further in View of Streit

Claims 157 and 267 are patentable over Kortge as discussed above. Streit, like Kortge, teaches a neural network and, thus, does not cure the deficiencies of Kortge. The combination of two neural networks cannot teach or suggest claims 158, 159, 268 and 269. Accordingly, the Section 103 rejection over the combination of Kortge and Streit should be withdrawn.

G. Claims 279-280, 289, 292-293, 302-303 and 321-322 Are Not Obvious Over Caid as Applied to Claims 271, 281, 291, 299 and 320 Above, and Further in View of Streit

The claimed invention distinguishes over conventional transforms and neural networks according to Caid as discussed above. Streit also teaches a neural network and, thus, does not cure the deficiencies of Caid. The combination of two neural networks cannot teach or suggest claims 279-280, 289, 292-293, 302-303 and 321-322. Accordingly, the Section 103 rejection over the combination of Caid and Streit should be withdrawn.

H. Claims 156, 270, 273, 275 and 284 Are Not Obvious Over Caid in View of Dickhaus

Conventional transforms and neural networks according to Caid do not teach or suggest the claimed invention as discussed above. The deficiencies of Caid are not corrected by Dickhaus, which teaches a similar approach to that of Caid involving formation of feature vectors from wavelet transformation of the input. On page 103, Dickhaus discloses that "[s]tatistical pattern recognition uses measurements and transforms of the pattern structure as feature vectors. Feature selection is often performed by sequential approaches, or <u>sometimes more or less intuitively by the experience of experts</u>." Dickhaus teaches an ad hoc approach. The combination of two neural networks cannot teach or suggest claims 156, 270, 273, 275 and 284. Accordingly, the Section 103 rejection over the combination of Caid and Dichaus should be withdrawn.

I. Response to Examiner's Argument that the Summary is Not Concise

The Examiner states on page 3 of his Answer that "(2) the summary is just repetition of the selected independent claims and is not concise." This argument is wholly without merit.

Appellant once again points out that that MPEP § 1206, "APPEAL BRIEF CONTENT," (5) "Summary of Invention," page 1200-8, states that "it is preferable to read the appealed claims on the specification and any drawing." On page 3-4 of the April 25, 2002 Appeal Brief, Appellant clearly stated that "[t]o facilitate ease of reading, only the independent claims have been discussed in this section with reference to the Figures. A complete copy of each appealed claim and a reading of each claim on the specification is set forth in Exhibit 5." Appellant spent a considerable amount of time providing specific references as to how each and every claim read on the specification and drawings in accordance with the MPEP. Appellant did not merely recite the independent claims as alleged by the Examiner, but went far beyond the minimum requirement of providing a mere summary. It is improper of the Examiner to now admonish Appellant for trying to help the Board of Appeals. Indeed, see page 1200-8 of the MPEP, which states "[w]hile reference to page and line number of the specification may require somewhat more detail than simply summarizing the invention, it is considered important to enable the Board to more quickly determine where the claimed subject matter is described in the application."

J. Claims 307-322 Fully Comply With 35 U.S.C. § 101

On pages 23-24 of the Examiner's Answer, the Examiner argues that:

> The data structure recited in Claims 307-322 comprises a collection of data objects. Claim referred in the case of In re Lowry, 32 USPTQ 2d 1031, 1035 (Fed. Cir. 1994) (now US patent 5,664,177) recites data structure. Because the claim dictates how application programs manage information, Lowry's Claim 1 defines *functional* characteristics of the memory (32 USPQ 2d 1034.) The *functional* characteristics make it patentable. However, the data structure of Claims 307-322 in the pending application does not control any computer program. The data structure is results of a collection of recognized patterns and their probabilities. Although the data structure is not static, it is passive data. It does not have any *functional* characteristics. The "activation of a data object" can be broadly interpreted just as "reading a data from a memory." The data structure is similar to the data compiled for a banking-account

> software. When a button on a computer screen is selected, data are read (activated) and displayed.
>
> In summary, the data structure recited in Claims 307-322 is passive in nature. It does not provide active function to change the software. Therefore, the invention as defined by Claims 307-322 is a non-functional data structure per se and therefore is non-statutory.

Appellant submits that claims 307-322 are not pure data as alleged by the Examiner. The activation of a data object results in "an efficient recognition of a pattern in newly presented information comprising data and input context."

The Examiner's abandonment of numerous objections and rejections raised to the specification and claims in his Answer in indicative of the unfair prosecution on the merits to which Appellant has been subjected. Despite the review and indication of allowability by six different Examiners, this case was suspiciously transferred to a seventh Examiner, whose lack of qualifications resulted in a specious rejection of Appellant's claims from which a first appeal was taken. [See Appellant's Brief dated April 3, 2001] Rather than respond to that first appeal, the PTO intensified its persecution of Appellant by transferring his application to the eighth and present Examiner. Incredibly, the rejections that followed were even more extreme, relying primarily on prior art technology that Appellant's specification expressly teaches against using, namely far removed neural networks. This latest rejection under Section 101 further demonstrates that extremism.

Conclusion

In view of the arguments presented hereinabove and in Appellant's Briefs on Appeal, all of the pending claims 51-322 fully comply with 35 U.S.C. §§ 101, 102, 103 and 112. Accordingly, Appellant respectfully requests that the Board withdraw the Examiner's rejections of claims 51-322 and allow all claims.

Respectfully submitted,

By ____________________

Jeffrey S. Melcher
Reg. No.: 35,950
Tel. No.: 202.261.1045
Fax. No.: 202.887.0336

Customer No. 20736

Manelli, Denison & Selter, PLLC
2000 M Street, N.W.
7th Floor
Washington, D.C. 20036-3307